US EPA ARCHIVE DOCUMENT

US EPA ARCHIVE DOCUMENT

# Guide to Data Management for the Environmental Results Program

**December 31, 2004**

***Prepared for:*** **USEPA National Center for Environmental Innovation**

***Prepared by:*** **Bill Clarke (Bill Clarke Consulting) and Michael Crow (The Cadmus Group, Inc.), with review assistance from Kathryn Waters (independent consultant).**

**Acknowledgments:**

The authors extend their gratitude to the following individuals who provided feedback on various iterations of this document:

- Kimberly Chesser (Delaware Department of Natural Resources and Environmental Control),
- Michael Redig (Florida Department of Environmental Protection),
- Robert Hawkins (Georgia Department of Natural Resources, Environmental Protection Division),
- Michelle Pruett (Tennessee Department Of Environmental Conservation),
- Michael Hurley (Massachusetts Department of Environmental Protection),
- Tara Velazquez (formerly of Massachusetts Department of Environmental Protection),
- Ronald Gagnon (Rhode Island Department of Environmental Management),
- Richard Enander (Rhode Island Department of Environmental Management),
- Thomas Armstrong (Rhode Island Department of Environmental Management),
- Scott Bowles (USEPA),
- Gregory Ondich (USEPA), and
- Marc Olender (USEPA).

US EPA ARCHIVE DOCUMENT

US EPA ARCHIVE DOCUMENT

US EPA ARCHIVE DOCUMENT

US EPA ARCHIVE DOCUMENT

US EPA ARCHIVE DOCUMENT

# INTRODUCTION

To many people, data management is little more than simply building a database and designing reports, or something for programmers and information technology (IT) people. As you will discover in this guide, sound data management is at the core of effective Environmental Results Program (ERP) implementation. It involves gathering, storing, analyzing and using your data to support the mission of your ERP.

Data management enables analysis and reporting of ERP inspection and certification data, as well as all of the data related to facility identification. It also establishes links between ERP data and existing warehouses of data on facility characteristics and performance. Well-planned data management will help lower the cost of supporting, growing, and managing the ERP. It will also make it possible to better use the data collected by the program.

## Purpose of the Guide

This guide is designed to familiarize you with the most important aspects of ERP data management, including data collection, analysis, and monitoring strategies. Think of it as a checklist for planning the data management of your ERP and finding ways to reduce manual (non-automated) operations. ERP is typically just one of a number of programs competing for resources. By planning ahead, you can automate many parts of the program, greatly reducing the cost and resources needed to support the program while potentially expanding your compliance monitoring efforts.

The guide includes lessons learned and suggested best practices from states that have already implemented ERP data management strategies. These are woven into the discussion as illustrations and examples of how good data management can enhance the effectiveness and viability of ERP.

Above all, the guide stresses best practices to achieve cost savings, efficiencies and data integrity through well-planned data management.

## Intended Audience

This guide will be most useful to ERP managers and their IT staff in states that are in the early stages of planning for ERP, or that are considering implementing new data management approaches. This document assumes that readers are already familiar with the general ERP approach and terminology.[1]

# ORGANIZATION OF THE GUIDE

This Data Management Guide is divided into three main parts, plus two appendices:

## Part 1: Program Design ("Business Needs") Questions

This section covers ERP policy and program design questions that help you figure out what you want your ERP to do. (Data management professionals may refer to these program design issues as "business needs.") These should be considered before developing a data management plan, to ensure your data management strategy suits your policy goals and processes. The questions in this section address

[1] Readers wishing to learn more about the basics of ERP may wish to review the EPA *ERP User's Guide for Government Agencies*. Readers wishing to learn more about statistical analysis of ERP data should review EPA's *A Generic Guide to Statistical Aspects of Developing an Environmental Results Program.* Both documents are available on EPA's ERP Web site.

program policies and procedures that will have a direct impact on how you plan and implement your approach to data management now and in the future. Some of these policies are fundamental to the mission of the program, while others are tactical procedures that will shape the day-to-day operation of the program. Part 1 specifically addresses data management issues related to the following topics:

- Requirements and incentives offered to facilities for ERP participation,
- Stakeholder communications,
- Identifying the universe of facilities,
- Random and targeted inspections,
- Delivery and receipt of certification forms,
- Changes to ERP over time, and
- Measurement indicators and reports.

Under each topic, several subtopics are presented. The discussion for each subtopic begins with one or more questions to help you think about data management in the context of your specific ERP. Those questions are followed by background information, where relevant.

## Part 2: System Requirements

The policy and procedural issues covered in Part 1 serve as the framework for the data management plan. Whereas Part 1 asks, "What do you want to do?" Part 2 asks, "How do you want to do it?" This section will define the requirements for your technology strategy. (For ease of reading, some issues related to system requirements issues are presented in Part 1.) Part 2 provides guidance on the following topics, to help you define your ERP's system requirements:

- Developing flowcharts for all of your ERP processes (such as certifications, inspections, etc.),
- Designing your database,
- Ensuring data integrity, and
- Defining necessary data reports and identifying automation opportunities.

## Part 3: Technology Strategy

This section asks the question: "What software and hardware will you need to support the program?" There are four considerations here: the cost of the technology, the ease of setting up and maintaining the system, the level of integration with existing data, and designing the system to meet future needs as the ERP grows and changes over time. Part 3 will help you:

- Identify the right strategy for your ERP, based on rules of thumb related to program size,
- Make technology investment decisions with the biggest payoff, and
- Understand key factors to consider in evaluating technology investments.

## Appendix 1: Top Issues to Remember

This section helps you change your focus from the details to the big picture questions related to ERP data management.

## Appendix 2: Project Plan Checklist

This section should help you develop a comprehensive data management plan.



US EPA ARCHIVE DOCUMENT

US EPA ARCHIVE DOCUMENT

# PART 1 – PROGRAM DESIGN ("BUSINESS NEEDS") QUESTIONS

The policies and procedures you put in place as you start your program will determine your approach to data management. Some questions are fundamental to the program and will have important implications for your approach to how you manage data. This section will describe some of these ERP design and implementation issues and how they will impact data management.

Note: When this guide refers to certifications, it is referring to all forms submitted by a facility in accordance with the policies of the program. This includes self-certification forms, non-applicability forms and return-to-compliance forms.

## Requirements and Incentives for Facilities to Participate in ERP

The types of requirements and incentives you choose for your ERP will influence the data that need to be tracked and potentially affect the implementation schedule and other aspects of your data management plan.

**Enforcing missed deadlines**

Will your agency want the option to take enforcement actions if facilities fail to submit self-certifications in a specified time period?

If so, you will need to track when self-certification forms are sent to facilities, when responses are received by the agency and when renewals are required. Most programs track the postal date for receiving certifications and other ERP compliance related communications.

Will your agency allow extensions of those time frames for extenuating circumstances (e.g., states of emergency, technical difficulties, mail strikes, etc)?

**Fees and penalties**

Will your program involve payment of new or existing fees or monetary penalties?

If so, this information will need to be tracked. Be sure to consider the following questions as you plan for data tracking:

- Who pays, how much do they pay and will the agency accept a single payment for multiple certifications in the event of a single owner of multiple regulated facilities?
- How are payments made and matched to the facilities?
- If your agency does not receive the fees and payments, who does? How and when do you determine who has paid and who has not?
- How about late fees, overpayments, refunds and underpayments? How and when are these reported and related back to the ERP data in your data management system? Are public hearings or other approval periods required for implementing new fees or for replacing existing fees?

US EPA ARCHIVE DOCUMENT

**Regulatory changes**

Are the data you will be collecting associated with other local, state or federal guidelines?

If so, how will you manage changes to these guidelines over time? For localities, how you will you manage variability across localities within your jurisdiction?

The data collected as part of an inspection may change from year to year to stay current with local, state or federal guidelines. So may the questions on the certification form. The data collection system should be flexible enough to accommodate these changes. Keep in mind that frequent changes will make it more difficult to compare yearly performance over time and therefore difficult to assess the overall performance of your program.

**Mandatory programs**

Will a new agency directive be needed to make participation mandatory? Will the forms need to be officially adopted by a rule making body? Is there a chance that the statutory or regulatory process could alter the data elements that need to be collected and thereby impact the database development process?

**Voluntary programs**

Will your program have voluntary certification?

If so, you may need to track data such as:

- When each facility joined,
- Whether and when it received particular incentives (e.g., recognition, reduced priority for inspections),
- Whether a given facility should be inspected to confirm reported compliance, and
- Whether each facility met the application deadline.

## Stakeholder Communication

Maintaining the support of the industry sectors, the public and other stakeholders is critical to the success of your ERP. This can be done through maintaining communications with stakeholders.

**External review**

Who will need to review your plans for ERP, and how long will the review process take?

Build into your schedule time for other agencies and departments, as well as any other key stakeholders, to review and approve your plans. One start-up ERP program reported that it took a year for it to get approval for their plan for ERP fees and penalties, adversely impacting the timeline for other activities, particularly data management.

US EPA ARCHIVE DOCUMENT

**Stakeholder contact data**

Do you need to stay in communication with a large number of stakeholders?

If so, tracking contact information through a stakeholder database or an automated email list may be useful. You may want to look at inexpensive commercial products for maintaining contact information or to assist in the cleaning and preparation of your existing data for use in mailings.

**Publicly available data**

Is it necessary, constructive, or practical to consider making facility data available to the general public on a Web site?

If so, consider the following questions:

- Would the publicly available data include owner names and results of certifications and enforcement actions?
- Would it be better to provide information about facilities individually or in aggregate?
- Would publicizing information about individual facilities reveal confidential business practices?
- Is it appropriate and fair to post the data if the data are possibly inaccurate or out of date?
- Would it raise homeland security concerns?

**Advisory committee**

Is it feasible to form an advisory committee drawn from the facilities or trade associations in the sector your ERP is targeting?

The advisory committee can provide immediate feedback on documents, guides and ERP policies, and help you maintain contact with the industry sectors over time. A committee can also be a good sounding board on potential data management issues that affect the targeted sector (e.g., publicly available data, Web-based data entry, etc.).

## Identifying the Universe of ERP Facilities

At the heart of your data management strategy is a database that will contain the identities of all facilities and all the information ERP generates about them. Developing and keeping this database up-to-date will be an on-going activity.

**Identifying ERP facilities**

What are the sources of lists of facilities (current agency databases, other state licensing agencies, professional associations, phone books, etc.) and how are they formatted (are they paper, printed labels, or data files)? Is it possible to include historical facility performance information in building

US EPA ARCHIVE DOCUMENT

the facility database? If there are several sources of facility data, how will you remove duplicate records?

Depending on the quality of available regulatory databases, you may have to build your initial database from scratch using other sources, such as state licensing agencies, professional associations, and phone books. This can be a time-consuming process that includes contacting facilities to verify if they are still in business and items such as owner's name, current address, and phone number. Allow enough time in your project plan for this activity. If you have the budget, you can hire temps or outsource this project to a telemarketing firm that specializes in these projects. Alternately, you might be able to use volunteer interns from a local educational institution. EPA's *Generic Guide to Developing Statistical Aspects of an Environmental Results Program* offers more advice on how to develop your list of facilities, also known as the facility universe.

The effort spent here can dramatically reduce the number of returned mailings and non-responders. It is advisable to track the sources of data for each facility record, so you can analyze which sources of data were most reliable after inspections are complete.

**Interface with other programs**

Will your data management system have to interface with other systems and regulatory programs? If so, consider the following questions:

- What are policies for updating facility data in situations where an ERP facility is under the jurisdiction of multiple environmental programs?
- How will you coordinate data and systems with these other programs?
- Is any coordination necessary with non-environmental governmental programs, outside the jurisdiction of your agency?
- How long will it take to get any approvals you need from inside and outside your agency?
- When you design your database, do you want to provide for multi-media information, with specific information for specific media and scalability for future growth of the program?

**Unique facility identifiers**

How will you index and uniquely identify facilities? Will the unique facility ID in your database need to be cross-referenced with other databases?

Check with your IT organization about how to uniquely identify your ERP facilities with a facility ID. Ask your IT organization

about how you can cross-reference your facilities database with other programs or data sources. It is critical to identify an indexing strategy that will allow you to easily cross-reference other databases before you implement your database.

**Tracking actors**

Are you tracking the practices of the business owner or the practices of the business?

Some programs are moving to what is called an "actor model" – they track data by the individual submitting the certification rather than by facility. If the person submitting the certification closes a business and opens a new business in the same or another ERP industry, this can be tracked.

The rationale is that business owners who ignore environmental best practices in one business may continue to ignore environmental best practices in other businesses in which they are involved. Decide early on whether you want to utilize an "actor model." Your choice will influence how you design your data collection and data management strategy. For instance, you may wish to require facilities to submit information on their corporate structure, to ensure that bad actors at the highest levels are not hidden behind local managers.

**Database update**

How do you plan to regularly update your database?

You will need to update your database of facilities after inspections and the certification period to correct any records that you found to be in error. If Web deployment is under consideration, then email addresses should be gathered as part of the facility information. Collecting email addresses will also enable the program to send out reminders to submit certification forms or Return-to-Compliance forms.

## Random and Targeted Inspections

ERP typically includes both random and targeted inspections, and data from these inspections form the backbone of the ERP's measurement system.

**Inspection reporting cycle**

How will inspection reports be initiated, filed, stored, retrieved and distributed?

Will they be paper, electronic or both? Do you need sampling and quality control reports? Do you need to communicate with other agency programs regarding pending inspections, to avoid duplicate effort? Do other programs need to be notified about ERP inspection results?



US EPA ARCHIVE DOCUMENT

US EPA ARCHIVE DOCUMENT

You might consider drawing a flowchart of the inspection process as it relates to ERP. Start with the initial decision to schedule an inspection and include any follow-up activities. (Similar issues related to certification forms are covered in the sections below.)

**Targeted inspections**

Will you have targeted as well as random inspections?

ERP data can enable an agency to effectively target its resources at the highest priority problems. If your agency conducts targeted inspections, consider how you will prioritize the targeted inspections.

- Will this involve a "scoring method"-- perhaps based upon Environmental Business Practice Indicators (EBPIs) -- that will need to be computed and stored with the facility's ERP record?
- Will it need to be stored anywhere else in the agency's data management system?

**Mobile computing for inspectors**

What are the pros and cons of equipping inspectors with laptops, tablet PCs or PDAs (personal digital assistants)? Can mobile computing investments provide value beyond ERP?

As discussed later, in Part 3, improving the efficiency of professional personnel, such as inspectors, should be seriously considered as part of any data-management strategy. Relieving inspectors of substantial data-entry responsibilities offers the potential to help them conduct more inspections or other professional duties. Automatic data-quality checks can also help ensure the soundness of data collected in the field. Further, wireless access to program data can also make inspectors more efficient in the field.

Care should be taken, however, because each mobile computing option may have many hidden costs: programming, maintenance, training, loss and theft. Equipping inspectors with laptops, tablet PCs or PDAs typically represents a department-wide strategic decision based on the value of mobile computing both to ERP and to other activities. If your ERP will be using scanning technology for certifications, you may want to consider using scanning for inspector checklists.

## Delivering the Certification Form to ERP Facilities

Most programs rely on paper certifications but a number are moving toward using the Web. While there is a large up-front cost, the Web improves the accuracy of your data and lowers the cost of data entry. The

nature of the selected sector will need to be considered when making delivery decisions. For example, a sector which relies very little on technology is probably not a good candidate for internet forms delivery and submission.

**Internet form delivery**

Will you make use of the Internet in the delivery process?

If you do use the Internet for forms distribution, consider the following questions:

- Can facilities download certification forms from your Web site?
- Will the targeted facilities know their identification numbers used by the data system to match the facilities to their historical data?
- Can they get their facility identification number from your mailings?
- How will you determine the form delivery date if the facility can download the forms off the Internet at any time it wishes?
- What format will you use to create and redisplay the forms?

Today most banks, insurance companies, mortgage brokers want documents that can be printed on a wide range of printing devices without distortion, and use a file format like Adobe's PDF. Word processing files often format and print the same document differently on different printing devices. This variability can create substantial difficulties for the agency in entering data received from facilities, particularly if it is using scanning devices. PDF files provide the added benefit of making it more difficult for a user to modify the form. Please note that software to create PDF documents is available from a number of vendors.

**Pre-printing facility data on forms**

Will you have sufficient data to pre-print facility information on certification forms?

When mailing out certification forms some programs print two labels: one for the outside envelope and another that is placed on the certification forms. Some programs believe that placing the label on the certification form increases the response rate. Preprinting the facility's agency identification on the form or label also facilitates data entry upon receipt of the certifications.

Anecdotal evidence indicates preprinting may not always be desirable, however. One ERP state avoids pre-populating paper forms, having found that new owners of an existing facility will tend to simply fill out the blank parts of the form, and not notify the agency of the change in ownership or other changes. For online submissions, this agency does pre-populate the form with existing data, but requires that the submitter provide a tax

US EPA ARCHIVE DOCUMENT

US EPA ARCHIVE DOCUMENT

| | |
|---|---|
| | ID number in order to complete the ERP online submissions. This requires new owners to declare the change in ownership. Such an approach is not as effective with paper submissions, since paper forms can be submitted incorrectly (e.g., without a tax ID number). Rectifying such problems can be an arduous task for an agency. |
| **Updating profile information on the certification form** | Have you given the facility an opportunity to update its profile information (e.g., owner, address, phone, and email) on the certification form?<br><br>ERP manages small businesses that are constantly changing, so it is important to take advantage of any opportunity to update information on the facility. Including an explicit change of address, phone, owner or email on the certification will improve the quality of your data. |
| **Tracking due dates** | How will you establish a due date?<br><br>If receipt of the certification form is a trigger for establishing a due date, then tracking the mail routing and receipt is critical.<br><br>Is the due date established from the date the certification form was mailed by you to the facility or by the date it was received by the facility? Or is the due date a standard regulatory date (e.g., always June 30)? |
| **Reminders for facilities** | Do you want to send out reminders to facilities to submit their certification forms, particularly if facilities have a long period of time to fill out the forms (e.g., a few months)?<br><br>If so, you may want to consider mail-merge software that runs against your master file of facilities. The ability to send out reminders cheaply is another incentive to collect e-mail addresses. |

## Receiving Certifications

The receipt and processing of certification forms can represent the most resource-intensive activity of the ERP, but may also present the best opportunity for automation and efficiency. ERP planners must consider a variety of different technical and resource issues in developing their strategies for certification receipt.

| | |
|---|---|
| **Single or multiple submission dates** | Do you have the staff capacity to send and subsequently receive all of your forms at one time or will you need to spread the |

US EPA ARCHIVE DOCUMENT

| | |
|---|---|
| | mailings out over time based on some internal data field (e.g. zip code)?<br><br>Spreading out the mailings over time is called "level-loading." Level-loading the mailings can make manual delivery and receipt more manageable relative to automated delivery and receipt, but agencies should recognize the trade-off that using an extended time period for level-loading can also substantially complicate the data analysis task. |
| **Tracking submission date** | How will you determine whether a facility submitted its certification form on time? By the postmark, or by the receipt date? Do both dates need to be tracked? What date will be used if the submitted certification is incomplete and is rejected and a corrected version resubmitted? |
| **Submission options** | How can ERP facilities file their certifications? Can they fax them in, mail them in, attach them to an email, or enter them through a Web form? Are the certification forms short enough to be entered through the phone? How will a facility keep a record of its submission, including the official answers it provided on the form?<br><br>Your answers to these questions will have a direct impact on how you design your data management system as will the variety of acceptable submission methods allowed.<br><br>If Web submissions are desired, then some level of identity authentication may be required, especially if the ERP hopes to use the information to take enforcement actions. Is there already a policy or system in place for e-submissions or will one need to be defined? |
| **Increasing Web certification** | If you are concerned that ERP facilities will not use the Web to certify, have you looked at strategies to increase Web certification?<br><br>The Web is increasingly becoming a part of life for businesses of all sizes. You should consider taking full advantage of it.<br><br>If you can generate sufficient numbers of Web certifiers, Web certification should lower your costs and increase the quality of certification data you collect from your target industries. Just using the Web to distribute your certification forms will lower your mailing costs dramatically. Instead of sending an entire certification questionnaire, you can send a postcard or one-page letter directing facility owners and operators to download forms from your Website. If your state has an eGovenment initiative, you should consider enabling forms to the filled out online. This will eliminate costly data entry from paper forms. |

US EPA ARCHIVE DOCUMENT

Many early-stage programs worry that their response rate will drop if facilities are "forced" to use the Internet. They fear that facilities are small businesses and will not have access to the Internet. However, some estimates (e.g., Nielsen, August 2004) indicate that more than 68 percent the US population now uses the Internet, and the experience of agencies that have implemented ERP programs suggests that this concern can be overcome.

For instance, one agency implementing a mature ERP recently sent a letter advising facilities to download paper forms from the Internet or to certify online. The letter mentioned that forms could be mailed out on request. The state found that over 60 percent of facilities in that sector provided certification information online, with most of the rest downloading forms and other documentation before submitting paper versions.

Before proceeding, however, you should carefully consider whether you believe you can attract enough facilities to certify on the Web to justify the investment in creating a Web certification system. This issue is particularly important for a voluntary program, when the exact number of facilities even participating in ERP is unknown at the outset.

**Entering certification data**

How will certification data be entered into the data management system?

In many instances, data from paper submissions will be keyed directly from a mailed-in form, but there are alternatives, particularly for high volumes. If you expect high volumes of forms, then a process involving scanning and automated data entry should be considered. (Both scanning and automated data entry are discussed later in this section.) As discussed above, Web submissions are also an option.

Automated data entry allows the scanning system to "read" the data on a scanned form. There are three types of automated data entry: optical character recognition (OCR) reads printed text, intelligent character recognition (ICR) reads hand-printed characters, and Mark Sense reads filled-in circles or checked boxes. If scanning is a consideration now or in the future, then it is important to work with the document imaging team to ensure the forms can be designed from the outset to maximize the benefits of OCR, ICR or Mark Sense.

**Signature issues**

If more than one party will need to sign the certification form (e.g., both owners and operators), how will an online system allow for double signatures?

In some sectors, variable owner/operator responsibilities could lead a state to require certification signatures from both owners and operators.

If each party fills out different parts of the form, determined by the parties themselves, can your online system accommodate this flexibility? How will your database record the different variations?

**Handling signatures on the Web**

If facilities enter certifications on the Web, how will you handle the requirement for a signature on the form?

As the Web becomes a larger part of the way we do business today, government and businesses have developed a number of ways to handle requirements for pen-on-paper (or "wet") signatures. Digital signatures typically involve downloading a form that users sign in order to apply for a digital signature. Users are assigned ID numbers they can enter on the Web form that serves as the signature authorization. Some programs collect the "wet" signature authorization on the first certification submitted by the facility. After that, they can use the Web forms for all future certifications.

The typical Web form application creates an on-line account with the user ID and password. Users sign-on to that account and download a form that is printed out, signed and sent in. A digital signature number is created that the user retrieves a few days later the next time they log into the site using their ID and password.

Other schemes do not require any paper signature. Some states allow users to "sign" on-line submissions by clicking a button on the Web site that affirms they are legally responsible for the submission. Check with your legal counsel to see what is acceptable in your state.

**Editing and validating data**

How will you edit and validate your data?

Most data-entry software provides methods for validating data as it is entered into the system. (This includes, for example, checking that dates have the proper format, as do phone numbers and zip codes.)

What is your policy for certifications with invalid data? Correct it and enter it anyway? Send it back? Call the facility on the phone and correct it as the data is being entered? Will it be necessary to grant extensions to the due dates?

Will it be appropriate to lock data entered by the facility from internal modification? Will it be necessary to track the identity



US EPA ARCHIVE DOCUMENT

| | |
|---|---|
| | of the person entering the data and the person quality checking and approving that entry? |
| **Incomplete certifications** | What constitutes an incomplete certification form? If specific questions are not answered? If there is no signature? Or name?<br><br>Should incomplete certification forms be entered into the data management system or not? If not, will you need to track which facilities submitted incomplete forms? If Web entry is allowed, should the incomplete entry be saved or rejected?<br><br>Will incomplete forms submitted on time be considered delinquent if the changes are not resubmitted within the initial deadline? Is your policy consistent for paper and electronic submissions? |
| **Non-responders** | Should non-responders be sent reminders?<br><br>If so, these may have to be identified in the database and mailings generated automatically. |
| **Non-applicable facilities** | To what extent do you want to track data for facilities that self-identify themselves as non-applicable (i.e., that declare that ERP does not apply to them)?<br><br>Do you want to be able to double-check these facilities, or have their records in the database in case their status changes or in case other programs within the agency find their data relevant? Do you want to be able to have information to support an enforcement case later?<br><br>Even if ERP truly does not apply to the facility, do you want to keep the data to provide a better picture of an individual who owns multiple facilities? This can be important if expansion of your ERP to other sectors or to a more comprehensive level within the current sector is planned.<br><br>Do you want to be able to link non-applicable facilities back to the database that originally provided their names, to better understand the accuracy of data sources? |
| **Web security** | If the data are entered or displayed on the Web, what level of identity validation and security will be needed? Does your agency already have relevant Web security protocols established? |
| **Scanning and automated data entry** | Should you implement scanning and automated data entry to process data from paper forms? |

US EPA ARCHIVE DOCUMENT

US EPA ARCHIVE DOCUMENT

As programs grow larger and add more sectors, they may need to store and manage more paper-based forms if Web-based data entry does not handle the bulk of the data volume.

Scanning allows forms and certifications to be stored as electronic images and managed as part of a document imaging and workflow process. This has the advantage of making forms easily accessible and available.

Automated data entry is the process of translating scanned digital images into usable data. There are three types of automated data entry: optical character recognition (OCR), intelligent character recognition (ICR), and Mark Sense. OCR software translates typewritten or printed text, whereas ICR software translates hand-written text. Mark Sense reads filled-in circles or checked boxes, such as for "yes/no" answers. OCR and ICR are necessary for handling non-standard information on forms, such as names, addresses and return-to-compliance plans. Because it deals with hand-written text, the accuracy of ICR is typically lower than that of OCR, particularly when dealing with atypical words (such as names).

Agencies should consider the pros and cons of scanning/automated data entry before choosing this approach.

*Pros*

- Eliminates paper files and makes documents easily available.
- Automated OCR/ICR may speed data entry.
- Makes copies of original documents with signatures available to inspectors.
- Makes diagrams and other graphical exhibits easier to share and manage.

*Cons*

- High initial costs for equipment and software can be difficult for a single program to justify.
- Automated data entry does not eliminate the need for data entry staff. People still need to scan the forms and verify that the data have been entered correctly. This can sometimes be time-consuming, and may nullify the efficiencies of scanning versus manual data entry.
- Forms will have to be re-designed to allow OCR/ICR.
- In the years to come the Web will be used more and more for eGovernment initiatives, making it difficult to get a good return on investment on scanning for ERP.
- Data quality control is not as robust as with Web-based data entry.

All programs have their own objectives and constraints. For some programs, paper forms may have a good long-term cost/benefit justification. But you should look carefully at your current and projected volumes and balance these against the expense of implementing and maintaining a scanning/imaging system for ERP. In the long term, most ERPs will likely be migrating to Web data entry. If other programs are using scanning and imaging, your ERP may be able to share the expense.

**Archiving**

Will you, by law, need to keep paper copies or printouts of submissions, or will electronic copies be acceptable? Is there already a protocol for archiving documents electronically?

If electronic copies are acceptable, be sure to account for the decrease in paper storage costs as part of the benefit of the project.

## Managing Changes Over Time

Managing ERP data presents a continuous challenge, and it is important to design your strategy to make it easy to maintain the data over time, and to accommodate changes to your program when necessary.

**Facilities opening and closing**

How will you handle facilities that open and close?

ERP deals with small businesses that are constantly in flux. Each year new businesses open, while others go out of business or change addresses, ownership or management. The following are important questions to consider:

- What are policies for handling changes in addresses, ownership or management?
- Do you update the existing record in the database or treat it as a "new facility" with a new facility number?
- Do you want to keep histories of facilities if they change names, owners or addresses?
- What are your policies for renewing and updating your list of ERP facilities from third-party sources?

**Facility updates online**

Will facilities be able to update their own facility profiles on-line?

Programs that use Web-based reporting are looking at ways they can allow facilities to update their own facility profile. This would lower costs and reduce staff time needed for routine maintenance. For programs with fees, changes to a facility's



US EPA ARCHIVE DOCUMENT

profile can mean a change in fees – for example, changes that involve a facility's size or capacity.

At least one ERP state has been reluctant to allow facilities to update their own profiles automatically over the Web out of concern that facilities would artificially and temporarily change their status to avoid paying fees. Instead, the state allows online requests for updates, but those requests must be approved by program staff.

**Changes in forms**

How frequently will the certification and/or inspection questions change? Is your database flexible enough to accommodate changes in questions and forms over time? How will the certifications be compared over time?

Each method for entering data into the system requires maintenance. If you change or update your certification, the system for entering data will also need to be updated. Be sure to allow enough lead time so these changes can be scheduled and implemented by your IT organization. If multiple methods for entering data are supported, then changes to the compliance forms or inspection forms will have multiple impacts.

## Measurement Indicators and Reports

While tending to the many details of ERP, it is critical not to forget your overarching goals, and to ensure that your data management system allows you to readily review the data in the format most meaningful to you.

**Measurement indicators**

How will you measure your success? What metrics will you use to evaluate the program? How will you track this data in your data management system over time?

**Reports**

What reports will you need – bar charts, measurements, etc? How will you analyze data and produce reports? Do you want to automate the reporting? Do you have access to the statistical expertise needed to appropriately characterize the program findings?

US EPA ARCHIVE DOCUMENT

US EPA ARCHIVE DOCUMENT

# PART 2 – SYSTEM REQUIREMENTS – HOW IS YOUR PROGRAM GOING TO WORK?

After you have developed the policies and procedures for your program, you are ready to design your data management system. Start by defining the tasks that will be performed on a day-to-day basis and how they fit together to form a coherent work flow. This will help you define what you need to automate. This section covers how to develop the requirements for your technology strategy. It is designed to help you understand the use of flowcharts, as well as issues related to database design, data integrity, reports, and automation.

## Flow Charts

Answers to the questions in Part 1 should help you design the processes of your ERP, allowing you to develop a "flow chart" of the program. Flow charts can be an extremely useful way to design your program and data management procedures. Flow chart diagrams have boxes and arrows that show how work flows through a business process. They allow you to walk through how you will process information and handle exceptions (errors or cases that deserve special attention). Flow charts are also a useful way to brainstorm procedural questions and identify potential problems before they become issues. In addition, they are excellent foundations for technical solution planning.

Your flow chart should help you clearly identify:

- **Steps in your process.** What is each step of your process, and who is responsible for each step?
- **Requirements for your system.** What are the data sources, actors, features and capabilities necessary to support your system? The flow chart will be the foundation from which to communicate clearly with your IT organization (or outside contractor) who will implement the system.
- **Data volume.** What volume of data (including both the number of forms and the amount of data on each form) will you need to process at peak times of the year?
- **Current databases.** What databases do you have in place on which you need to build?
- **Other infrastructure.** With what other IT systems will you need to interface?
- **Timing.** Are there specific timing constraints on how the data is processed?
- **Universe size.** How many facilities are within the current and foreseeable ERP universe? How will you quantify volumes and work loads?
- **Reports/analysis.** What reports and analysis will you need?

Figure 1 (below) presents a sample ERP flow chart for the certification process. You would also want to develop flow charts covering other data management activities, such as those related to the random inspection process. The discussion following Figure 1 explains how the flow chart works to help you better understand the concept and how it might apply to your ERP.

US EPA ARCHIVE DOCUMENT

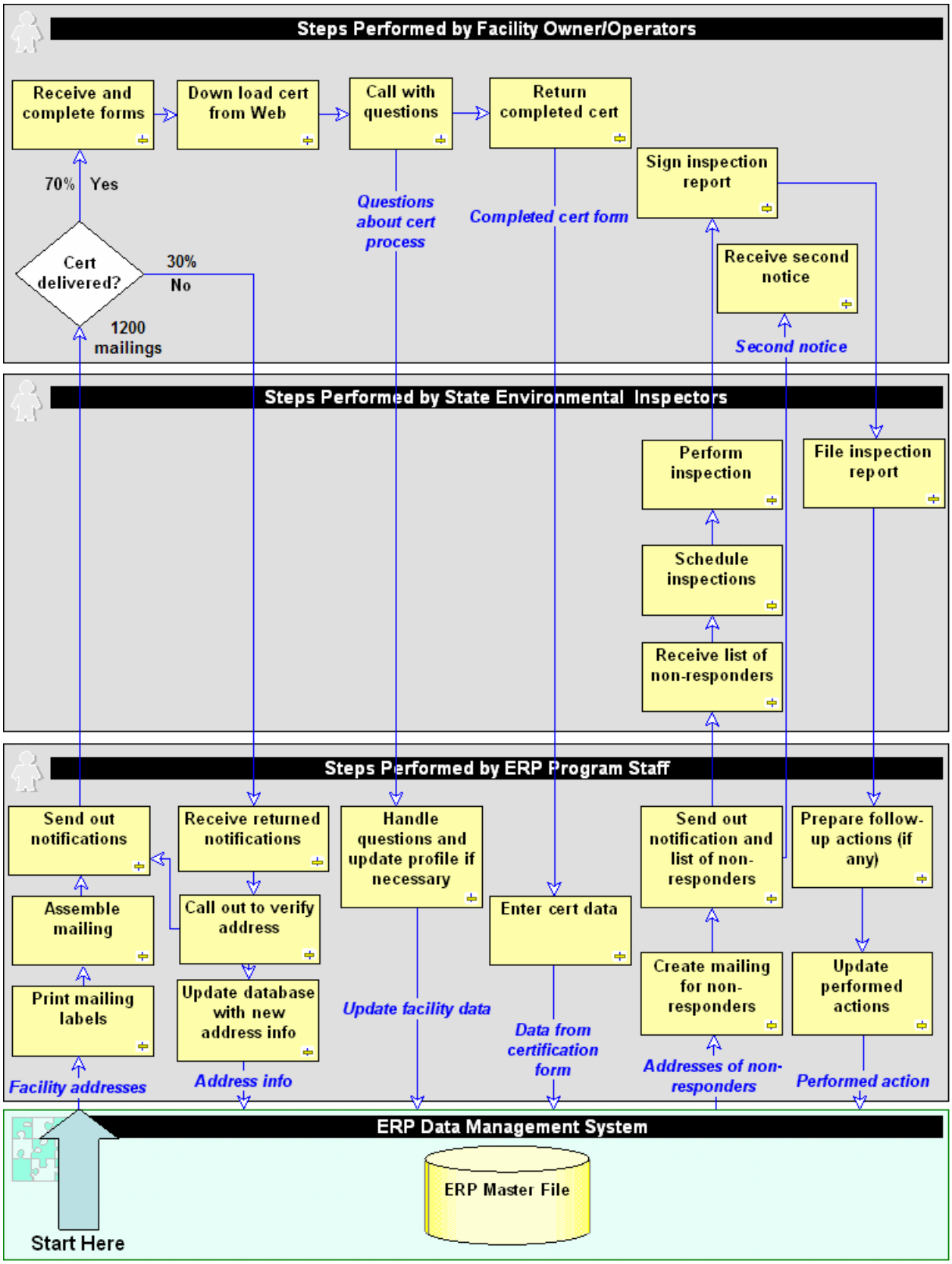



**Figure 1: Sample ERP Flow Chart (Certification Process)**

**Description of flow chart**

The horizontal bands in Figure 1 are called "swim lanes" and identify organization units or actors that perform steps in the process. Swim lanes can also show organizations outside your department (in this case facility owner/operators) or systems (in this case the ERP Data Management System that manages the ERP Master file).

The boxes are the steps in the operations or tasks performed by each organization unit. The arrows connecting the boxes are "messages." These represent the data that flow between the steps in the process. They can be reports or notifications (e.g., phone calls and emails).

Notice that the flow chart can also be annotated with volume data or the probability that a step will be taken. For instance, this flow chart notes that the agency anticipates 1200 mailings of certification forms, and expects that 30 percent will be returned as undeliverable.

**Using the flow chart to develop system requirements**

Every time a message goes into or out of the ERP Data Management System, the system must be able to handle that event. This shows a clear requirement for the IT organization that will be implementing the system. In this example, the system needs to be able to:

- Generate mailing labels for mailers (message "Facility addresses")
- Indicate that an address is invalid (message "Invalid address flag")
- Allow sector managers to update facility data (message "Update facility data")
- Enter data from certification forms (message "Data from certification form")
- Generate mailing labels for non-responders (message "Non-responders")
- Be able to update performed actions on a facility (message "Performed actions")

**Charting your way to success**

Develop flow charts for each interaction with the facility, such as a return-to-compliance plan submitted (or not submitted), facility closed or out of business, or filing a non-applicable form. Include in the flow charts interactions or processes supported by other systems.

For example, if your program is mandatory or involves fees, another department could handle payments. How will your ERP process interface with that department's system?

US EPA ARCHIVE DOCUMENT

There is no need to fit everything on one diagram. Include a narrative with the flow chart that gives a high level description of each step.

**Building your organization**

Notice also that the flow chart shows who is supposed to do what, and in what order. It can also show when and where you need reports – the messages between steps can be reports. You will find flow charts are a valuable training tool for new employees.

## Designing Your Database

After you have designed your ERP process, you can start on what you will need in your database.

**Make an initial list of data items**

When you start developing the requirements for your database, develop an initial list of everything you think you want to keep in it (e.g., the facility's name and address, the contact name, certification data). Use a flow chart to be sure you have included everything you need.

If you are aware of additional data items you may want to collect in the future, be sure to communicate this to your database designers so that the database will be able to accommodate these items.

**Use the flow chart**

To check yourself, look at the messages on the flow chart that go into and out of the system. Identify all the data items that might be needed in each message and how the ERP process uses that data.

For examples, see Figure 1:

- The message "Facility addresses" is used to print mailing labels. It should contain the mailing address of the facility. Is this different from the owner's address? Do you need space for two addresses?
- If a mailing is returned, you will need a way to flag that an address is invalid. You may also want to mark the date when the mailing was returned.
- The process shows a second mailing to non-responders. The system will need a way to identify who did not respond by a certain date.

As you work through what your own ERP process will need, the requirements for your database will become more focused.



US EPA ARCHIVE DOCUMENT

| | |
|---|---|
| **Create logical groupings** | Cluster the data items into logical groupings. Data that are used together or relate to a common subject should be grouped together. Each grouping should have a unique ID. IDs need not be arbitrary numbers. For instance, they can consist of a combination (or concatenation) of several data items.<br><br>For example, one grouping will surely be the data from a certification form. Assuming you do not want to store multiple copies of the same form submitted on the same day, forms could have an ID that includes the facility ID, form type and date. Which date? Most likely the submittal date. |

## Data Integrity

In any system managing substantial amounts of data, such as ERP, mistakes can be made. Developing a plan for maintaining data integrity can help minimize mistakes and ensure that your database closely resembles the information provided by inspectors and facilities.

| | |
|---|---|
| **Data glossary** | Have you included a task for documenting your data?<br><br>A data glossary (or data dictionary) documents the meaning and use of each data item in your database and ERP system. It should include the source of the data, how and when it is updated, and how it used. This is critical for consistent use of the data. |
| **Granting update privileges** | Who can update what data and when?<br><br>"Updating" includes creating, changing and deleting data in your system. Whenever a staff member changes any data in your ERP system, you should have an audit trail showing who changed what, and when. |
| **System of record** | Do you have a single database that is the authority or system of record?<br><br>Avoid situations where you must manually update a data item in multiple locations (e.g., update a change of address in several different files). This invites errors and inconsistencies. One system should be the "authority" or system of record. Updates to this system should be propagated to other systems. |
| **Control and audit reports** | Have you provided for control reports?<br><br>Whenever a data item is updated, you should consider having a record of when and who updated the data item. |
| **Error logging and monitoring** | Can you verify that all modifications to data are being performed as expected? |

US EPA ARCHIVE DOCUMENT

| | |
|---|---|
| | Systems that rely on automated updating events or include cascading updates need to ensure there is error logging. Those logs should be checked daily by the Data Administrator or the individual assigned to Quality Assurance. |
| **Backup and restoration** | Can you restore corrupted data or data that was changed erroneously (deleted, updated, or created)?<br><br>If backup and restoration is handled by your IT department, verify its cycles and determine whether it meets your needs. For example, if a backup is made weekly, could you lose data if you have to restore mid-week? |
| **Tracking updates** | Can you track updates?<br><br>Most database management systems (DBMS) have the capability to track changes to the data in the database. Make sure your program takes full advantage of these capabilities to track when a data item is updated, created or deleted. |
| **Accommodate changes in policies** | Can your database support changes to policies and procedures?<br><br>If you include start- and end-dates for data and relationships, you can build in features that can accommodate changes in policies. |
| **Data administration** | Is there someone responsible for data administration?<br><br>The data administrator is responsible for data integrity and data usage, including report design. For example, among your "non-responders" will be facilities that have closed, gone out of business or moved. The data administrator will ensure that updates to the ERP database are valid and correct. The data administrator will also work with your IT department to ensure that changes to the design or usage of the data are consistent with the policies of your ERP. As you expand to other sectors, your data administrator will help design the database to ensure that changes in the design do not compromise existing data. It is important to assign this role to a particular person, to make sure it gets done. |

US EPA ARCHIVE DOCUMENT

## Identifying Needed Reports and Opportunities for Automation

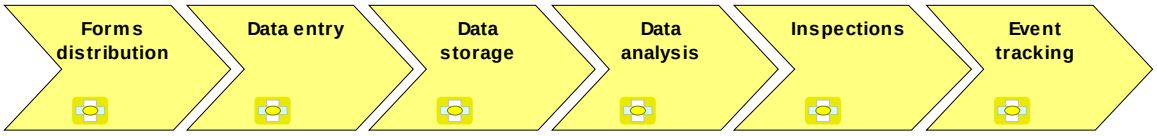
The flow charts will yield a number of requirements for implementing your data management system. Requirements are what you will need to give to programmer or Information Technology (IT) staff to implement your system. IT staff will also be very interested in the flow charts to understand what the system needs to do and what data it needs to store, manage, and analyze. The flow charts will also help you identify reports that you need and opportunities for you to automate.

| | |
|---|---|
| **What reports will you need?** | Reporting requirements will change with your program. Reports should give you the metrics you need to manage the program. |
| **How are you being evaluated?** | It is useful to look at each major function in your process and ask how each person in that job function will be evaluated in terms of performance. What information does each role need to optimize its performance? What reports does each role need to measure and improve performance? |
| **How is the program being evaluated?** | Look at the overall ERP program and define the metrics (measures) on which it will be evaluated. What reports do you need to be able to measure, evaluate and improve on that performance? |
| **Automate the bottlenecks** | The flow charts can help you identify where you should be looking to apply automation to improve performance and lower costs. |

US EPA ARCHIVE DOCUMENT

# PART 3 – TECHNOLOGY STRATEGY – AUTOMATION OPTIONS

This section will help you identify technology options for different-sized programs and understand how to make sound investments. Regardless of the size or process flow of your program, from a technology point of view, there are six activities you will have to support. These are:




**Forms distribution** – how do you distribute certification forms, workbooks and instructions to your ERP sectors?

**Data entry** – how do you enter the certification data received from facilities into your system?

**Data storage** – how do you store and integrate the data you need to manage?

**Data analysis** – how do you evaluate the data you have collected?

**Inspections** – how do you support data collection and reporting for inspections?

**Event tracking** – how do you track performed actions, outstanding issues and follow-up activities?

How you handle each of these activities depends on how large your program is today and how large your program will be in the future. When you think about the size of your program, you need to think about the volume of data you will need to manage. The factors that influence ERP data volume include the following:

- Number of facilities per sector,
- Number of data items per facility,
- Number of sectors,
- Number of facilities being inspected,
- Number of certification submissions (particularly for voluntary programs), and
- Frequency of data submissions (e.g., certifications and inspections).

## Find the Right Strategy for Your Program

ERPs can grow quickly, especially when the concept shows success. This section will give you a feel for what technology strategies may be appropriate for different-sized programs.

The size of an ERP can generally be described by the total number of facilities across all sectors and objects that fall under the mission of ERP. For the purposes of this analysis, we will use the cut-off points below. Please note that programs with a significant volume of data per facility may find that data management strategies for programs with larger numbers of facilities will be appropriate. As noted above, data volume is influenced by several factors.

US EPA ARCHIVE DOCUMENT

- **Small programs** – fewer than 1,000 facilities in all sectors starting from year one. As the program moves into its third or fourth year, it should consider data-management strategies recommended for medium sized programs.

- **Medium-sized programs** – 1,000 to 5,000 facilities across all sectors starting from year one. As the program moves into its third or fourth year, it should consider data management strategies recommended for large programs.

- **Large programs** – more than 5,000 facilities across all sectors starting from year one. The technology strategy selected should support all future growth and data storage requirements.

## Small Programs

Programs with fewer than 1,000 facilities in all sectors starting from year one.

| | |
|---|---|
| **Forms distribution** | Complete certification kits can be mailed out directly to facilities. Each kit includes certification forms and workbooks. Kits can also be made available at industry meetings and through your advisory council, if you have one. If your state agency has a Web site, kits can be posted on the site and downloaded by facilities. Your agency can direct facilities to the Web site by mailing out notifications or reminders. |
| **Data entry** | Manual data entry from paper forms, or if the agency has a defined Web entry process in place, direct Web entry may also be an option. |
| **Data storage** | Programs in the first year may want to start with desktop data storage applications, such as a spreadsheet or integrated database package. This will give you the appropriate level of support and the flexibility to change the structure of the database as requirements change.<br><br>In year two or three, as the amount of data you manage grows, you can "graduate" to desktop or server databases. Desktop databases are easier to manage but are designed for single user access. Server-based databases are multi-user but are much more expensive to set up and administer.<br><br>Regardless of the solution you select, it is important that you develop and follow strict daily backup procedures from day one. Also be sure to build your initial system with enough flexibility and consistency with your enterprise standards to minimize rework as the system evolves. |
| **Data analysis** | In year one, the native spreadsheet capabilities should be adequate to support your analysis. You should plan to move toward an integrated database in later years to handle the increased volume of data and more complex reporting. |

US EPA ARCHIVE DOCUMENT

**Inspections** Small programs may not require automated support for inspections. Inspection targets are selected as time and budget allow – particularly if inspectors are also supporting other programs. Inspection reports will typically be paper-based and filed on paper.

**Event tracking** Event tracking will be largely manual, involving email, phone messages and personal contact. The ERP database should be able to record "performed actions" and follow-up activities. Entry of these events needs to be quick and as accurate as possible to ensure the users will enter the required data.

## Medium-Sized Programs

Programs with 1,000 to 5,000 facilities across all sectors starting from year one.

**Forms distribution** Web-based distribution should be considered to reduce the expenses of postal mailing. Postcard mailers can be sent directing facility operators to download forms and workbooks from the Web site. For those facilities without Internet access, mailings will still need to be an option.

**Data entry** The program should be considering Web-based data entry. Paper forms will still be part of the program, but efforts should be made to develop a Web forms entry. This will dramatically reduce the cost of data entry and increase the accuracy of the data. Web forms, where facilities log-in and enter data on-line, have the advantage of assisting the user in providing a complete and accurate submission. If an entry is wrong or out of range, the user can be directed to correct the problem.

Paper forms will still be part of the program, but the volume should not be large enough to justify the cost of scanning or automated data entry unless you have more that 3,000 facilities and no Web-based certification entry.

**Data storage** Evaluate a server-based database that can support enterprise applications. You will need to manage an increasing volume of data in years one, two, and three, and you will need the performance and scalability provided by an enterprise database management system.

This class of database is expensive to set-up, implement and maintain. Planning will save money in the long run because poor design decisions can be expensive to correct. Many of your other agency programs will probably already have server-based solutions in place with enough capacity to support your new ERP, without significant impact to performance.

**Data analysis** Enterprise database management systems typically have report writers available as an option. You should also consider third party add-on



US EPA ARCHIVE DOCUMENT

US EPA ARCHIVE DOCUMENT

| | |
|---|---|
| | software that is easier to use or has analysis features you need. Remember that automating reports can help make your most valuable human resources more efficient.<br><br>Note: Select report-writing software that has a large number of users in the marketplace, so it will be easier to find contractors and consultants to implement reports. |
| **Inspections** | Medium-sized programs may begin to look at providing inspectors with laptops, tablet PC's or personal digital assistants (PDAs) with electronic inspection report forms included as an application. This has the advantage of reducing paperwork and increasing the productivity of inspectors, particularly if they can reference background information on facilities. The decision to provide inspectors with mobile computers is typically a department-wide decision that rests on the value of the investment both in ERP and other activities. |
| **Event tracking** | Event tracking should be built into the database design from day one. This includes tracking activities and scheduled events such as inspections, non-compliance notifications, reminders, and direct contacts with facilities. If audits of changes are also desired, the level of auditing, the events and information must also be noted. |

## Large Programs

More than 5,000 facilities across all sectors starting from year one.

| | |
|---|---|
| **Forms distribution** | Web-based distribution should be considered very seriously to reduce the expenses of mailing, if a substantial number of facilities can be reasonably expected to provide data electronically. Postcard mailers should be sent directing facilities to download forms and workbooks from the Web site. |
| **Data entry** | As much as possible, the program should be using Web-based data entry. Paper forms should be discouraged.<br><br>If it is not possible to use the Web for direct entry and the program must use paper forms, consider scanning and automated data entry (OCR or ICR). Up-front expenses are large, but without automated data entry it may not be possible to control the paper flow. Several ERPs are moving toward enterprise document management systems. It is possible to leverage those enterprise systems for your scanning with minor enhancements. |
| **Data storage** | Use an enterprise database hosted on an enterprise server. Data volumes will grow quickly and the program will need be able to manage the volume of data. Consider hiring or retaining a professional database programmer or administrator. |

| | |
|---|---|
| **Data analysis** | Reporting requirements will be complex and require a third-party report writer package. |
| **Inspections** | Large programs may have already provided inspectors with laptops, tablet PCs or even personal digital assistants (PDAs). To make inspectors even more productive, consider providing them with access to data and documents in the field. This can be done by allowing them to log into a secure Web site where they can download documents and upload reports and data. |
| **Event tracking** | Event tracking should be built into the database design from day one. This includes tracking activities and scheduled events such as inspections, non-compliance notifications, reminders, and direct contacts with facilities. |

## Invest Your Money Wisely

Regardless of your program's size, you should be realistic about how long it will take to recover the costs of your technology investment. A cost-benefit analysis for ERP automation will typically show a trade-off between the costs of doing a task manually (the cost of human resources) versus the costs of automation (the investment in technology). An incremental investment in technology can result in an incremental decrease in the cost for human resources for a task (and people can be deployed elsewhere). At some point, however, this balance breaks down and costs of automating a task are more than doing the task manually.

You will get a higher return on your investment in technology if you can improve the efficiency of your more valuable human resources assets. For most ERPs, the most valuable (operational) human resources assets are sector managers and inspectors. You will get a higher return on your investment in technology by making these people more efficient.

For example, sector managers and inspectors spend many unproductive hours sifting through data to identify issues that require action. Perhaps a facility's certification form has inconsistencies that merit follow-up. Finding all facilities with noteworthy inconsistencies could require your professional staff to create and run repeated database queries, month after month, year after year. Your ERP may see substantial gains in efficiency by translating those activities into automatic queries of the database. An advanced data management system can use automatic queries to identify cases for follow-up and notify the professional staff responsible for acting on the items. The system could even generate reminders if such items go unattended.

Inspectors can also spend substantial amounts of time transcribing data from paper forms into electronic format, and verifying the accuracy of the data. As discussed in Part 1, mobile computing solutions for inspectors (such as laptops or PDAs) can help focus their efforts on tasks that merit their professional attention.

When your automation strategy is aimed at minimizing the repetitive work conducted by professional staff, you will get a higher return on your investment and the entire program wins.

US EPA ARCHIVE DOCUMENT

## Consider All the Costs of a Technology Investment

Technology investments have four components that should be factored into their total cost, or what is called the "Total Cost of Ownership"(TCO):

1. **Acquisition cost** – cost for the initial purchase and support contract;
2. **Implementation cost** – cost for customization and adapting a technology to your program;
3. **Deployment costs** – costs for training, cut-over (transitioning operations to the new system) and hiring, training and deploying staff;
4. **Maintenance** – yearly operational costs for operating the system (staffing, support and routine enhancements).

The total cost of ownership is used to compute your return on investment (ROI) over a period of time. ROI is typically your net savings (savings minus your TOC) over a period of time expressed as a percentage of your investment (or TOC) over that same period of time. The typical time horizon for IT investments is three to five years and the typical target ROI is usually around 15 to 20 percent. Speak to your financial department about how your department treats ROI and its criteria for acceptable ROI .

Work with your IT department and other internal stakeholders to develop a sound TCO estimate, as well as an understanding about who is paying for what. Be sure to discuss variables that might lead to cost escalation and how any escalations in cost will be paid for. You may wish to record the understanding in writing.

# CONCLUSION

Based on your reading of the preceding sections, you should now feel familiar with all aspects of ERP data management, including data collection, analysis, and monitoring strategies. You should be better prepared to plan the data management of your ERP and find ways to increase automation, where feasible. You should also feel well-equipped to think through the feasibility of different data management strategies. As you move forward in planning and implementation, you may wish to refer back to different parts of this guide; the detailed table of contents is intended to help you remember where to find particular issues you are interested in revisiting. You may also want to read and use the two appendices: Appendix 1, which offers a reminder of eight of the most important issues to remember, and Appendix 2, which provides a checklist of the many details to consider in developing a data management project plan.

US EPA ARCHIVE DOCUMENT

US EPA ARCHIVE DOCUMENT

## APPENDIX 1: TOP ISSUES TO REMEMBER

The sections above are intended to illuminate the multitude of critical details involved in ERP data management. However, as you develop and implement plans for ERP data management, it is easy to become immersed in the details and forget important, "big picture" issues. The items below represent important issues that ERP planners should consider early on, and come back to occasionally. Doing so can help ensure that the data management planning and implementation goes more smoothly.

1. **Coordination.** With whom do you need to coordinate to implement ERP? How do their schedules and availability fit with your plans? Will your data management system have to interface with other systems and other regulatory programs? Who will need to review your plans, and how long will it take? Are there data management efforts underway within your agency or state that you can leverage or that you must be consistent with?

2. **Spillover benefits.** Are there opportunities for your technology investment to provide benefits beyond ERP, perhaps justifying greater automation?

3. **Online certification.** Does Web certification makes sense for your ERP, considering the data volume, the cost of setting up and maintaining the system, and the capabilities and characteristics of the facilities?

4. **Readiness for change.** Is your data management system scalable and flexible? That is, how easily will you be able to adapt your data management approach if circumstances change and/or your number of facilities grows?

5. **Maximize human resources.** How can you use better data-management approaches to maximize the use of your most valuable human resources?

6. **Uncertainty.** Is your ERP a pilot project whose fate depends on project outcomes? If so, have you taken this uncertainty into account in identifying the most cost-effective technology investments?

7. **Reports.** What are the most important outputs you want to see from your data management system? Is your system designed to easily provide you with these reports?

8. **Garbage in, garbage out.** Have you developed a plan for data integrity, to ensure that the information underlying your reports is accurate?

US EPA ARCHIVE DOCUMENT

## APPENDIX 2: PROJECT PLAN FOR DATA MANAGEMENT

In the process of considering the multitude of issues presented above, you may find it very helpful to record your considerations and decisions in writing. One good way to do this is to develop a data management project plan. Below is a checklist of tasks you may want to include in your project plan:

**Business Needs**

- Develop requirements for participating in ERP
- Identify key dates and deadlines for certification
- Finalize policies for fees and penalties
- Develop incentives for participation in voluntary programs
- Schedule policy review period (if required)
- Develop methods for maintaining stakeholder contacts
- Develop an advisory committee
- Investigate and evaluate data sources to develop initial list of facilities
- Develop a database of facilities
- Determine how your ERP will interface with other programs
- Develop inspection reporting cycle
- Investigate any regulatory changes that could affect your program
- Decide on how to deliver the certification forms to ERP facilities
- If you are using the Internet for forms delivery, coordinate with Webmaster of your Web site
- Design certification forms including method for pre-printing facility data on forms
- Decide on method for tracking key dates
- Determine how to handle the receipt of certifications
- Determine submission options
- Speak with your legal counsel about signature issues, particularly if your facilities are filing certifications on-line
- Work with your IT department to design and implement a data entry form for entering certification data
- Develop rules for editing and validating data for the data-entry forms and provide these rules to your IT department
- Decide how you want to handle incomplete certifications
- Decide how you want to handle non-responders
- Decide how you want to handle non-applicable facilities
- Decide how you want to handle facilities that have opened or closed
- Develop a method for scheduling targeted inspections
- Develop measurement indicators so you can measure your success
- Decide how you want to handle changes to your certification questions
- Design management and operation reports

**System Requirements – How is your program going to work?**

- Develop flow charts for how your program will work. You may want to develop flow charts for each of the following:
    - Printing and sending out certifications or announcements if you are making them available in the Web
    - Receiving and entering data into your data management system
    - Handling facility information changes
    - Handling returns and non-responders
    - Handling non-compliance
    - Handling inspections
    - Recording and tracking fees
    - Generating and distributing reports and results

- List the data items you want to track in your data management system
- Use the flow charts to validate your list
- Create logical groupings of the data items
- Share this list of data items with the person within your IT organization who will be implementing the data management system
- Review your flow charts with members of the IT organization and work with them to identify opportunities for automation and to ensure they understand the business process and needs
- Develop a plan for ensuring data integrity
- Define the reports needed and provide these specifications to your IT organization who will be implementing the data management system

**Technology Strategy – Automation Options**

- Work with your IT organization to develop a short-term and long-term technology strategy that is consistent with the IT standards and leverages other IT initiatives. Discuss with them:
    - Web data entry and security
    - Scanning/automated data entry
    - Laptops, tablet PCs or PDAs for inspectors
    - Upgrades and adding enhancements to the data management system
    - Total cost of ownership and return on investment requirements

US EPA ARCHIVE DOCUMENT